(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 241 662 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
18.09.2002 Bulletin 2002/38

(51) Int Cl.7: **G10L 15/14**, G10L 15/20

(21) Application number: **02100251.4**

(22) Date of filing: **14.03.2002**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **14.03.2001 US 275487 P**

(71) Applicant: **Texas Instruments Incorporated**
Dallas, Texas 75251 (US)

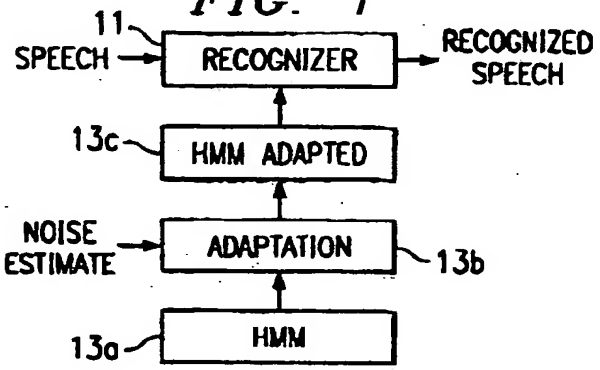
(72) Inventor: **Gong, Yifan**
Plano, Texas 75025 (US)

(74) Representative: **Holt, Michael**
Texas Instruments Ltd.,
EPD MS/13,
800 Pavilion Drive
Northampton Business Park,
Northampton NN4 7YL (GB)

(54) Method of speech recognition with compensation for both channel distortion and background noise

(57) A method of speech recognition with compensation is provided by modifying HMM models trained on clean speech with cepstral mean normalization. For each speech utterance the MFCC vector is calculated for the clean speech database. This mean MFCC is added to the original models. An estimate of the background noise is determined for a given speech utterance. The model mean vectors adapted to the noise are determined. The mean vector over the noisy speech space is determined and this is removed from the model mean vectors adapted to noise to get the target model.

FIG. 1



EP 1 241 662 A2

NSDOCID: <EP_____1241662A2_I_>

Description

Field of Invention

[0001] This invention relates to speech recognition and more particularly to compensation for both background noise and channel distortion.

Background of Invention

[0002] A speech recognizer trained with relatively a quiet office environment speech data and operating in a mobile environment may fail due to at least to the tow distortion sources of back ground noise and microphone changes. The background noise may be from a computer fan, car engine, and/or road noise. The microphone changer may be due to the quality of the microphone, hand-held or hands-free and, a position to the mouth. In mobile application of speech recognition, both the microphone conditioner and background noise are subject to change.

[0003] Cepstral Mean Normalization (CMN) removes utterance mean and is a simple and effective way of dealing with convolutive distortion such as telephone channel distortion. See "Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker
Identification and Verification" of B. Atal in Journal of Acoustics Society of America, Vol. 55: 1304-1312, 1974. Spectral Subtraction (SS) reduces background noise in the feature space. See article "Suppression of Acoustic Noise in Speech Using Spectral Subtraction" of S.F. Boll in IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-27 (2): 113-129, April 1979. Parallel Model Combination (PMC) gives an approximation of speech models in noisy condition from noise-free speech models and noise estimates. See "An Improved Approach to the Hidden Markov Model Decomposition of Speech and Noise" of M.J. F. Glaes and S. Young in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 1, pages 233-236, U.S.A., April 1992. The techniques do not require any training data.

[0004] Joint compensation of additional noise and convolutive noise can be achieved by the introduction of a channel model and a noise model. A spectral bias for additive noise and a cepstral bias for convolutive noise are introduced in an article by M. Afify, Y. Gong, and J. P. Haton. This article is entitled "A General Joint Additive and Convolutive Bias Compensation Approach Applied to Noisy Lombard Speech Recognition" in IEEE Trans. on Speech and Audio Processing, 6(6): 524-538, November 1998. The five biases can be calculated by application of Expectation Maximization (EM) in both spectral and convolutive domains. A procedure by J.L. Gauvain, et al, is presented to calculate convolutive component, which requires rescanning of training data. See J.L. Gauvain, L. Lamel, M. Adda-Decker, and D. Matrouf entitled "Developments in Continuous Speech Dictation using the ARPA NAB News Task." In Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, pages 73-76, Detroit, 1996. Solution of the convolutive component by steepest descent method has also been reported. See Y. Minami and S. Furui entitled "A Maximum Likelihood Procedure for a Universal Adaptation Method Based on HMM Composition." See Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, pages 129-132, Detroit, 1995. A method by Y. Minami and S. Furui needs additional universal speech models, and redestination of channel distortion with the universal models when channel changes. See Y. Minami and S. Furui entitled "Adaptation Method Based on HMM Composition and EM Algorithm" in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, pages 327-330, Atlanta 1996.

[0005] The techniques presented by M.F.J. Gales in "PMC for Speech Recognition in Additive and Convolutional Noise," Technical Report TR-154, CUED/F-INFENG, December 1993 needs two passes of the test utterance, e.g., parameter estimation followed by recognition, several transformations between cepstral and spectral domains, and a Gaussian mixture model for clean speech.

[0006] Alternatively, the nonlinear changes of both type of distortions can be approximated by linear equations, assuming that the changes are small. Jacobian approach, which models speech model parameter changes as the product of a jacobian matrix and the difference in noisy conditions, and statistical linear approximation are along this direction. See S. Sagayama, Y. Yamaguchi, and S. Takahashi entitled "Jacobian Adaptation of Noisy Speech Models," in Proceedings of IEEE Automatic Speech Recognition Workshop, pages 396-403, Santa Barbara, CA, USA, December 1997. IEEE Signal Processing Society. Also see "Statistical Linear Approximation for Environment Compensation" of N.S. Kim, IEEE Signal Processing Letters, 5(1): 8-10, January 1998.

[0007] Maximum Likelihood Linear Regression (MLLR) transforms HMM parameters to match the distortion factors. See "Maximum Likelihood Linear Regression for Speaker Adaptation of Continuous Density HMMs" by C.J. Leggetter and P.C. Woodland in Computer, Speech and Language, 9(2): 171-185, 1995. This method is effective for both sources but requires training data and introduces the dependence to the speakers.

SDOCID: <EP_____1241662A2_I_>

Summary of Invention

[0008] In accordance with one establishment of the present inventor a new method that Handles simultaneously noise and channel distortions to make a speaker independent system robust to a wide variety of noises and channel distortions.

Description of Drawing

[0009]

Figure 1 illustrates a speech recognizer according to one embodiment of the present invention; and

Figure 2 illustrates the method of the present invention generating

Description of Preferred Embodiments of the Present Inventions

[0010] Referring to Fig. 1 there is illustrated a speech recognizer according to the present invention. The speech is applied to recognizer 11. The speech is compared to Hidden Markov Models (HMM) 13 to recognize the text. The models initially provided on those with speech recorded in a quiet environment and the microphone of good quality. We want to develop a speech model set suitable for operating in the simultaneous presence of channel/microphone distortion and background noise. In accordance with the present invention, a speech model set is provided using statistics about the noise and speech. A low computation cost method integrates both PMC and CMN.

[0011] Referring to Figure 2, the first Step 1 is to start with HMM models trained on clean speech, with cepstral mean normalization. We modify these models to get models to compensate for channel/microphone distortion (convolutive distortion) and simultaneous background noise (additive distortion). For an HMM model, we have a lot of parameter but only change one subset of the parameters and that is mean vectors $m_{p,j,k}$. The mean vectors $m_{p,j,k}$ of the original model space is modified where p is the index of the Probability Density Function (PDF), j is the state and k is the mixing component.

[0012] The second Step 2 is to calculate which is the mean mel-scaled cesptrum coefficients (MFCC) vector over the trained database. Scan all data and calculate the mean to get $\hat{b}$.

[0013] The third Step 3 is to add mean $\hat{b}$ to each of this mean vector pool represented by $m_{p,j,k}$ equation (1) to get:

$$\bar{m}_{p,j,k} = m_{p,j,k} + b. \tag{1}$$

For example, there could be 100 PDFs, 3 states per PDF and 2 vectors per state, or a total of 600 vectors.

[0014] The fourth Step 4 is for a given input test utterance, an estimate of the background noise vector $\tilde{x}$ is calculated.

[0015] Let $u^l \triangleq [u_1^l, u_2^l, \cdots u_D^l]^T$ and $v^l \triangleq [v_1^l, v_2^l, \cdots v_D^l]^T$, where *l* means that the values are represented in log-spectral domain.

[0016] We introduce the combination operator $\oplus$ such that:

$$w^l \triangleq u^l \oplus v^l = [w1^l, w2^l, \cdots wD^l]^T \tag{2}$$

with

$$w_j^l = \log(\exp(u_j^l) + \exp(v_j^l)) \tag{3}$$

[0017] In Step 5, we calculate the mean vectors adapted to the noise $\tilde{x}$ using equation 4.

$$\hat{m}_{p,j,k} = IDFT(DFT(\bar{m}_{p,j,k}) \oplus DFT(\tilde{X})). \tag{4}$$

NSDOCID: <EP_____1241662A2_I_>

where *DFT* and *IDFT* are, respectively, the DFT and inverse DFT operation, $\overline{\mathbf{m}}_{p,j,k}$ is the noise compensated mean vector.

[0018] Equation 4 involves several operators. DFT is Discrete Fourier Transform and IDFT is the Inverse Discrete Fourier Transform. The ⊕ is an operation with two vectors. A ⊕ B = C. How ⊕ is defined, we look at equations 2 and 3. Equation 2 says the operation + operates on two vectors u and v and the result is a vector of D dimension or $[w_1, w_2, \ldots w_D]^T$ where T is the transposition. We take the two vectors and produce another vector. We need to specify each element in the resultant vector. Equation 3 says that the jth element in that vector (w_j) is defined by the exponential of the element of u added to the exponential if the jth element of v and take the log of the combination of the exponential of u added to the exponential of the j the element of v. This completes the definition of Equation 4.

[0019] In the following steps, we need to remove the mean vector $\hat{b}$ of the noisy data y over the noisy speech space N (from the resultant model). One may be able to synthesize enough noisy data from compensated models but this requires a lot of calculation. In accordance with the present invention the vector is calculated using statistics of the noisy models. The whole recognizer will operate with CMN (cepstral mean normalization mode), but the models in Equation 4 are no longer mean normalized. We have dealt with additive noise. The second half of the processing is removing the cepstral mean of our models defined in Equation 4. This is not difficult because we have the models in Equation 4. In Step 6, we need to integrate all the samples generated by Equation 4 to get the mean. Mean is $\hat{b}$. Equation 5 is this integration.

[0020] Let , be the variable denoting PDF index, *J* be the variable for state index, and K be the variable for mixing component index.

$$\hat{b} = E\{y\}$$

$$= N^{y}\Sigma_p\Sigma_j\Sigma_k P,(p)\, P_{J|,(j,p)}\, P_{K|,,J}(k|p,j)\, p_{Y,,,J,K}(y|p,j,k)\, dy \qquad (5)$$

Since

$$p(y|p,j,k) = N(y, IDFT(DFT(\overline{\mathbf{m}}_{p,j,k}) \oplus DFT(\tilde{X})), E_{p,j,k}) \qquad (6)$$

We have

$$\hat{b} = \Sigma_p\Sigma_j\Sigma_k P,(p)P_{J|},(j|\,p)p_{K,,J}(k|p,j)\,\hat{m}_{p,j,k} \qquad (7)$$

Equation 7 shows that $\hat{b}$ can be worked out analytically, and it is not necessary to do the physically generation and integration. The final result is Equation 7 which is the integration into several sums. Sums over probability density functions and the sum over states and sum over mixing components. Then you have several quantities. The P_H is the probability of having the PDF index. The P_J given H is the probability of being in the state if given the PDFp. The next is the probability the mixing component $p_{\hat{}}j$ given we have the PDF index. The mean vector of the compensated mode. To make this complete we remove this $\hat{b}$ from the compensated model to get the target model. This is Step 7. The target model is:

$$\dot{m}_{p,j,k} = \hat{m}_{p,j,k} - \hat{b} \qquad (8)$$

This is what we want to load into our recognizer. This operation is done for each utterance.

[0021] Calculation of $\hat{b}$ thus requires the knowledge of probability of each PDF. There are two issues with P (, = *p*) :

- It needs additional storage space.
- It is dependent of the recognition task e.g. vocabulary, grammar.

[0022] Although it is possible to obtain that probability, we want to consider the following simplified cases.

[0023] This operation to calculate the $\hat{b}$ can be simplified with three approximations. The first one uses equal probabilities for P_H (*p*) or constraint C.

;DOCID: <EP_____1241662A2_I_>

1. Use equal probabilities for $P_H(p)$:

$$P_H(p) = C \tag{9}$$

2. Use equal probabilities for $P_H(p)$, P_J I, $P(j|p)$ and $P_{K|.,.J}(k|h,j)$.

$$P_H(p) = C$$

$$P_{J|H}(j|p) = D \tag{10}$$

$$P_{K|H,.J}(k|p,j) = E$$

3. In fact, the case described in Eq-10 consists in averaging the compensated mean vectors $\overline{\mathbf{m}}_{p,j,k}$. Referring to Eq-4 and Eq-1, it can be expected that the averaging reduces the speech part $m_{p,j,k}$ just as CMN does. Therefore, Eq-7 could be further simplified into:

$$\hat{b} = IDFT\,(DFT(b) \oplus DFT\,(\tilde{X}))\,. \tag{11}$$

The model $m^T_{p,j,k,t}$ is then used with CMN on noisy speech. Unfortunately, $\hat{b}$ is a function of both channel and background noise in all above cases. In other words, in presence of noise, there is no guarantee that the channel will be removed by such a vector, as is for CMN.

[0024] A subset of WAVES database containing hands-free recordings was used, which consists of three recording sessions: parkedtrn (car parked, engine off), parked (car parked, engine off), and city-driving (car driven on a stop and go basis).

[0025] In each session, 20 speakers (10 male) read 40 sentences each, giving 800 utterances. Each sentence is either 10, 7 or 4 digit sequence, with equal probabilities. The database is sampled at 8kHz, with MFCC analysis frame rate of 20ms. Feature vector consists of 10 statis and 10 dynamic coefficients.

[0026] HMMs used in all experiments are trained in TIDIGITS clean speech data. Utterance-based cepstral mean normalization is used. The HMMs contain 1957 mean vectors, and 270 diagonal variances. Evaluated on TIGIDIT test set, the recognizer gives 0.36% word error rate.

[0027] To improve the performance in noisy environment, the variances of the Gaussian PDFs can be MAP adapted to some slightly noisy data, e.g. WAVES parked_eval data. Such adaptation will not affect recognition of clean speech, but will reduce variance mismatch between HMMs the noisy speech. The new algorithm is referred to as JAC (joint compensation of additive noise and convolutive distortion).

Table 1:

word error rate (WER) (%) as function of driving conditions and compensations methods		
	PARKED	DRIVING
BASELINE	1.38	30.3
CMN	0.59	18.8
PMC	1.74	6.29
JAC FULL PROB (Eq-7)	0.84	1.93
JAC CONST PDF PROB (Eq- 9)	0.84	2.00
	1.16	2.18
JAC AVERAGE (Eq-10)	0.48	1.91
JAC SIMPLIFIED (Eq-11)		

MSDOCID: <EP 1241662A2 I >

[0028] Table-1shows that:

- Compared to noise-free recognition (WER) (0.36%), without any compensation (BASELENE) the recognition performance degrades severely.
- CMN effectively reduces the WER for parked data, but is not effective for driving conditions where additive noise becomes dominant.
- PMC substantially reduces the WER for driving conditions, but gives poor results for parked data where microphone mismatch is dominant.
- All JAC cases give lower WER than non-JAC methods.
- Simplifying Eq-7 to Eq-9 then to Eq-10 results in progressive increase in WER, although the degradation is not severe. Especially, information in PDF probability is not critical to the performance.
- Simplified JAC gives lowest WER in all tests. For this hands-free speech recognition, the new met reduces word error rate by 61% for parked condition and 94% for city driving condition.

Claims

1. A method of modifying HMM models trained on clean speech with cepstral mean normalization to provide models that compensate for simultaneous channel/microphone distortion and background noise (additive distortion) comprising the steps of:

   for each speech utterance calculating the mean mel-scaled cepstrum coefficients (MFCC) vector $\hat{b}$ over the clean database;
   adding the mean MFCC vector $\hat{b}$ to the mean vectors $m_{p,j,k}$ of the original HMM models where p is the index of PDF, j is the state, and k the mixing component to get in $m_{p,j,k}$;
   for a given speech utterance calculating an estimate of the background noise vector $\tilde{x}$;
   calculating the model mean vectors adapted to the noise $\tilde{x}$ using $\hat{m}_{p,j,k} = \text{IDFT}(\text{DFT}(\bar{m}_{p,j,k} \oplus \text{DFT}(\tilde{x}))$ to get the noise compensated mean vector where the Inverse Discrete Fourier Transform is taken sum of the Discrete Fourier Transform of the mean vectors $\bar{m}_{p,j,k}$ modified by the mean MFCC vector $\hat{b}$ added to the Discrete Fourier Transform of the estimated noise $\tilde{x}$; and
   calculating the mean vector $\hat{b}$ of the noisy data over the noisy speech space, and removing the mean vector $\hat{b}$ of the noisy data from the model mean vectors adapted to noise to get the target model.

2. The method of Claim 1 wherein the step of calculating the mean vector $\hat{b}$ of the noisy data over the noisy speech space will calculate the vector using statistics of noisy model using :

$$\hat{b} = \Sigma_p \Sigma_j \Sigma_k P_{,}(p) P_{J|,}(j|p)\, P_{K|,,J}(k|p,j)\, \hat{m}_{p,j,k}$$

   where *H* is the variable denoting PDF Index J is the variable for the sate index and K is the variable for mixing component index

3. The method of Claim 2 wherein said calculating the mean vector $\hat{b}$ uses equal probabilities for $P_H(p)$

$$p_H(p) = C.$$

4. The method of Claim 2 wherein equal probabilities for $P_H(p)$, $P_{J|,}P(j|p)$ and $P_{K|,,J}(k|h,j)$ *is* used.

$$P_H(p) = C$$

ISDOCID: <EP_____1241662A2_I_>

$$P_{J|H}(j|p) = D$$

$$P_{K|H,J}(k|p,j) = E$$

5. The method of Claim 3 wherein mean vector $\hat{b}$ becomes equal to:

$$\hat{b} = IDFT\ (DFT(b) \oplus DFT\ (\tilde{X}))\ .$$

6. A method of speech recognition with compensation for channel distortion and background noise comprising the steps of:

providing HMM models trained on clean speech with cepstral mean normalization; and modifying said models by the method of any preceeding claim

7. A speech recognizer with compensation for channel distortion and background noise comprising in combination:

adapted HMM models generated by modifying HMM models trained on clean speech with cepstral mean normalization wherein said models are adapted by:

for each utterance:

calculating the calculating the mean mel-scaled cepstrum coefficients (MFCC) vector $\hat{b}$ over the clean database;
adding the mean MFCC vector $\hat{b}$ to the mean vectors $m_{p,j,k}$ of the original HMM models where p is the index of PDF, j is the state, and k the mixing component to get in $m_{p,j,k}$;

for a given speech utterance calculating an estimate of the background noise vector $\tilde{x}$;

calculating the model mean vectors adapted to the noise $\tilde{x}$ using

$$\hat{m}_{p,j,k} = IDFT\ (DFT\ (\overline{m}_{p,j,k} \oplus DFT\ (\tilde{X}))$$

to get the noise compensated mean vector where the Inverse Discrete Fourier Transform is taken sum of the Discrete Fourier Transform of the mean vectors $\overline{m}_{p,j,k}$ modified by the mean MFCC vector $\hat{b}$ added to the Discrete Fourier Transform of the estimated noise $\tilde{x}$; and
calculating the mean vector $\hat{b}$ of the noisy data over the noisy speech space, and removing the mean vector $\hat{b}$ of the noisy data from the model mean vectors adapted to noise to get the adapted model;and means for comparing the adapted model to the speech input utterance to recognize the input speech

8. The recognizer of Claim 7 wherein the step of calculating the mean vector $\hat{b}$ of the noisy data over the noisy speech space will calculate the vector using statistics of noisy model using :

$$\hat{b} = N^{y}\Sigma_p\Sigma_j\Sigma_k P,(p)P_{J|},(j|p)P_{K|,,J}(k|p,j)\,\hat{m}_{p,j,k}$$

where *H* is the variable denoting PDF Index J is the variable for the sate index and K is the variable for mixing component index 6. The model of Claim 5 wherein the step of calculating the mean vector $\hat{b}$ of the noisy data over the noisy speech space will calculate the vector using statistics of noisy model using :

$$\hat{b} = \Sigma_p\Sigma_j\Sigma_k P,(p)\ P_{J|,}(j|p)P_{K|,,J}(k|p,j)\,\hat{m}_{p,j,k}$$

\NSDOCID: <EP_____1241662A2_I_>

where H is the variable denoting PDF Index J is the variable for the sate index and K is the variable for mixing component index

9. The recognizer of Claim 8 wherein said calculating the mean vector $\hat{b}$ uses equal probabilities for $P_H(p)$

$$P_H(p) = C.$$

10. The recognizer of Claim 8 wherein equal probabilities for $P_H(p)$, P_J I, P(j|p) and $P_{K|,,J}(k|h,j)$ *is* used.

$$P_H(p) = C$$

$$P_{J|H}(j|p) = D$$

$$P_{K|H,J}(k|p,j) = E$$

11. The method of Claim 8 wherein mean vector $\hat{b}$ becomes equal to:

$$\hat{\mathbf{b}} = IDFT\ (DFT(\mathbf{b}) \oplus DFT\ (\tilde{\mathbf{X}}))\ .$$

12. A method of speech recognition with simultaneous compensation for both channel/microphone distortion and background noise comprising the steps of:

modifying HMM models by a method as claimed in any of claims 1 to 5.

SDOCID: <EP_____1241662A2_I_>

FIG. 1

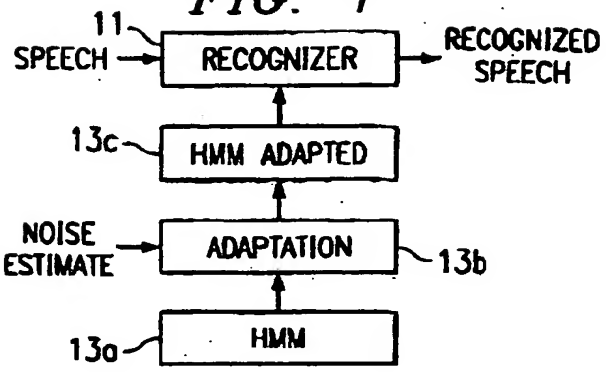
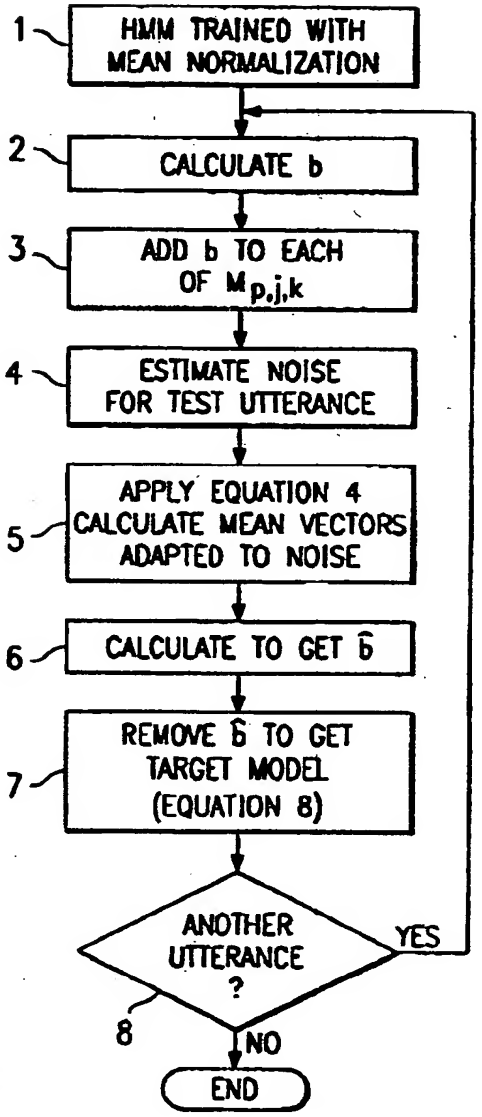


FIG. 2



NSDOCID: <EP_____1241662A2_I_>

THIS PAGE BLANK (USPTO)

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 241 662 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
18.02.2004 Bulletin 2004/08

(51) Int Cl.7: **G10L 15/20**

(43) Date of publication A2:
18.09.2002 Bulletin 2002/38

(21) Application number: **02100251.4**

(22) Date of filing: **14.03.2002**

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **14.03.2001 US 275487 P**

(71) Applicant: **Texas Instruments Incorporated**
Dallas, Texas 75251 (US)

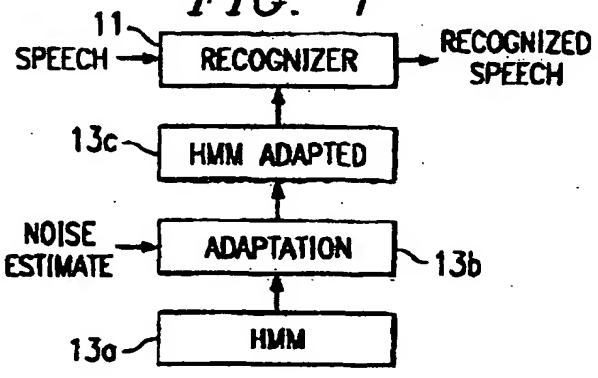
(72) Inventor: **Gong, Yifan**
Plano, Texas 75025 (US)

(74) Representative: **Holt, Michael**
Texas Instruments Ltd.,
EPD MS/13,
800 Pavilion Drive
Northampton Business Park,
Northampton NN4 7YL (GB)

(54) Method of speech recognition with compensation for both channel distortion and background noise

(57) A method of speech recognition with compensation is provided by modifying HMM models trained on clean speech with cepstral mean normalization. For each speech utterance the MFCC vector is calculated for the clean speech database. This mean MFCC is added to the original models. An estimate of the background noise is determined for a given speech utterance. The model mean vectors adapted to the noise are determined. The mean vector over the noisy speech space is determined and this is removed from the model mean vectors adapted to noise to get the target model.

FIG. 1



EP 1 241 662 A3

NSDOCID: <EP_____1241662A3_I_>

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 10 0251

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	YAMAMOTO H ET AL: "Fast speech recognition algorithm under noisy environment using modified CMS-PMC and improved IDMM+SQ" ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1997. ICASSP-97., 1997 IEEE INTERNATIONAL CONFERENCE ON MUNICH, GERMANY 21-24 APRIL 1997, LOS ALAMITOS, CA, USA,IEEE COMPUT. SOC, US, 21 April 1997 (1997-04-21), pages 847-850, XP010225927 ISBN: 0-8186-7919-0 * abstract; figure 1 * * page 848, left-hand column, line 15 - right-hand column, line 14 * ---	1-12	G10L15/20
A	SHOZAKA M ET AL: "A non-iterative model-adaptive E-CMN/PMC approach for speech recognition in car environments" 5TH EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY. EUROSPEECH '97. RHODES, GREECE, SEPT. 22 - 25, 1997, EUROPEAN CONFERENCE ON SPEECH COMMUNICATION AND TECHNOLOGY. (EUROSPEECH), GRENOBLE: ESCA, FR, vol. 1 OF 5, 22 September 1997 (1997-09-22), pages 287-290, XP002203281 * abstract; figure 3 * * page 289, right-hand column, line 1 - page 290, left-hand column, line 14 * --- -/--	1-12	TECHNICAL FIELDS SEARCHED (Int.Cl.7) G10L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	22 December 2003	Zimmermann, E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

SDOCID: <EP_____1241662A3_I_>

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 10 0251

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	KOSAKA T ET AL: "Instantaneous environment adaptation techniques based on fast PMC and MAP-CMS methods" ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 1998. PROCEEDINGS OF THE 1998 IEEE INTERNATIONAL CONFERENCE ON SEATTLE, WA, USA 12-15 MAY 1998, NEW YORK, NY, USA,IEEE, US, 12 May 1998 (1998-05-12), pages 789-792, XP010279271 ISBN: 0-7803-4428-6 * abstract; figure 1 * * page 791, left-hand column, line 20 - line 46 * -----	1-12	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	22 December 2003	Zimmermann, E

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

\SDOCID: <EP_____1241662A3_I_>

THIS PAGE BLANK (USPTO)